

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 7, 2005, 21:32:16 ; (without alignments)  
Sequence: 1 LRALVHGNLQYAEIPKSEI.....RRUDAFHAIYNDWRGENGEP 364  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqb, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA,\*

1: /cgn2\_6/prodata/1/1aa/5A-COMB-pep:,\*  
2: /cgn2\_6/prodata/1/1aa/5B-COMB-pep:,\*  
3: /cgn2\_6/prodata/1/1aa/6A-COMB-pep:,\*  
4: /cgn2\_6/prodata/1/1aa/6B-COMB-pep:,\*  
5: /cgn2\_6/prodata/1/1aa/PCRTUS-COMB-pep:,\*  
6: /cgn2\_6/prodata/1/1aa/backfile1.pep:,\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1879	100.0	364	4	US-10-166-606-4
2	1589	84.6	4	US-08-613-212A-4	
3	140.5	7.5	647	1	US-07-894-212A-8
4	140.5	7.5	649	1	US-07-894-212A-2
5	140.5	7.5	650	1	US-07-893-212A-1
6	105	5.6	227	4	US-09-107-532A-5554
7	101.5	5.4	360	4	US-09-710-2120-2150
8	101.5	5.4	376	4	US-09-710-2120-2106
9	101	5.4	329	1	US-08-270-013B-2
10	101	5.4	329	1	US-08-838-41B-2
11	97	5.2	653	4	US-09-540-3128
12	94.5	5.0	367	3	US-09-134-001C-3678
13	92	4.9	375	4	US-09-710-2120-1840
14	92	4.9	845	4	US-09-248-796A-5394
15	91.5	4.9	741	4	US-09-252-991A-31448
16	91.5	4.9	1009	4	US-09-693-446-4
17	91	4.8	1031	4	US-09-543-881A-8245
18	90.5	4.8	406	3	US-09-134-001C-3570
19	90.5	4.8	789	4	US-09-248-796A-19294
20	90	4.8	1171	4	US-09-248-796A-0522
21	90	4.8	3854	4	US-09-949-016-7876
22	89.5	4.8	322	4	US-09-448-039A-22383
23	89	4.8	857	4	US-09-248-796A-22741
24	88.5	4.7	399	4	US-09-252-991A-22741
25	88.5	4.7	764	3	US-09-235-451-36
26	88.5	4.7	4	US-09-248-796A-36	
27	88.5	4	993	4	US-09-894-998A-50

ALIGNMENTS

RESULT 1  
US-10-166-606-4

Sequence 4, Application US/10166606  
; PATENT NO. 6644756:  
; GENERAL INFORMATION:  
; APPLICANT: Murphy, Dennis  
; ATTORNEY: Reid, John  
; TITLE OF INVENTION: ALPHA GALACTOSIDASES AND METHODS FOR  
; TITLE OF INVENTION: MAKING AND USING THEM (Amended)  
; FILE REFERENCE: 0910-004005  
; CURRENT APPLICATION NUMBER: US/10/166, 606  
; CURRENT FILING DATE: 2003-01-31  
; PRIOR APPLICATION NUMBER: US 09/407, 806  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 08/613, 220  
; PRIOR FILING DATE: 1996-03-08  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 364  
; TYPE: PRT  
; ORGANISM: Thermococcus alcaliphilus  
US-10-166-606-4

Query Match 100.0%; Score 1879; DB 4; Length 364;  
Best Local Similarity 100.0%; Pred. No. 5.2e-191;  
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRALVHGNLQYAEIPKSEI-PKVIEKAYVPIVETIPLKEIPGPNLNTGYIKPLPKDID 60  
Db 1 LRALVHGNLQYAEIPKSEI-PKVIEKAYVPIVETIPLKEIPGPNLNTGYIKPLPKDID 60

QY 61 LVKGIGIASDLIEBIGTSYTHAILPLPLSLRSVEAQVQDRREVKEELFEVSPKGFWPLAY 120  
Db 61 LVKGIGIASDLIEBIGTSYTHAILPLPLSLRSVEAQVQDRREVKEELFEVSPKGFWPLAY 120

QY 121 DPILPAIKLKGKYEYIFADEGAMLSAHSIANSIAKPKIPIYPLIQAQRERFRYISVLLG 180  
Db 121 DPILPAIKLKGKYEYIFADEGAMLSAHSIANSIAKPKIPIYPLIQAQRERFRYISVLLG 180

QY 181 LRLRKGKLUVPEGKUTLKVAKDTEAVPVVAVNTAWMLGIGRLPLMPKKVASTEDKD 240  
Db 181 LRLRKGKLUVPEGKUTLKVAKDTEAVPVVAVNTAWMLGIGRLPLMPKKVASTEDKD 240

QY 301 DKSRLWREBGNALMUSYNMGEELAFLAENSNDARGMERPLPERLDAFRAYNDWRGE 360  
Db 301 DKSRLWREBGNALMUSYNMGEELAFLAENSNDARGMERPLPERLDAFRAYNDWRGE 360

RESULT 2  
US-08-613-220B-4  
; Sequence 4, Application US/08613220B  
; Patent No. 5956751  
; GENERAL INFORMATION:  
; APPLICANT: Murphy, Dennis  
; TITLE OF INVENTION: ALPHA-GALACTOSIDASE  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson, P.C.  
; STREET: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: US  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/613.220B  
; FILING DATE: 08-MAR-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Halle, Ph.D., Lisa A.  
; REGISTRATION NUMBER: 38,347  
; REFERENCE/DOCKET NUMBER: 09010/004001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-678-5070  
; TELEFAX: 619-68-5099  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 346 amino acids  
; TYPE: amino acid  
; TOPOLOGY: Linear  
; MOLECULE TYPE: Protein  
; FRAGMENT TYPE: Internal  
; US-08-613-220B-4

Query Match 84.6%; Score 1589; DB 2; Length 346;  
Best Local Similarity 95.1%; Pred. No. 3,3e-160;  
Matches 346; Conservative 0; Mismatches 0; Indels 18; Gaps 18;

QY 1 LRALVFGHNLQYABIKPSBIPKJIEKAVIPVETLKEIPBIPFGNITGTYKLKPQDID 60  
Db 1 LRALVFGHNLQYABIKPSBIPKJIEKAVIPVETLKEIPBIPFGNITGTYKLKPQDID 57  
Qy 61 LVKGGIASDLIELIGTSYTHAIPPLPLSRVEAQDREVKERBLFVEVKPGFWLPFLAY 120  
Db 58 LVKGGIASDLIELIGTSYTHAIPPLPLSRVEAQDREVKERBLFVEVKPGFWLPFLAY 114  
Qy 121 DPTIPAIKLDNGYELFADGRMLFSAHNSAIKPKIOLYPIKIIKAQKRRYIVSILG 180  
Db 115 DPTIPAIKLDNGYELFADGRMLFSAHNSAIKPKIOLYPIKIIKAQKRRYIVSILG 171  
Qy 181 LRLRKATKLVFEGKVTKAVKQTEAVPVWVVAINTAVMLGIGRLPLMPKKVAVSWEDK 240  
Db 172 LRLRKATKLVFEGKVTK-VKQIEAVPVWVVAINTAVMLGIGRLPLMPKKVAVSWEDK- 228  
Qy 241 NLLYLYGTDIEFFIGYRDIAGYRMSVEGLLRLVIDLSELCLPSBLKHSGRELYRLTSSWAP 300

RESULT 3  
US-07-894-212A-8  
; Sequence 8, Application US/07894212A  
; Patent No. 536883  
; GENERAL INFORMATION:  
; APPLICANT: ASADA, KIYozo  
; APPLICANT: UEMOJI, TAKASHI  
; APPLICANT: MUKAI, HIROYUKI  
; APPLICANT: KATO, IKUNOSHIN  
; APPLICANT: LADEMAN, KENNETH  
; APPLICANT: ANFINSEN, CHRISTIAN  
; TITLE OF INVENTION: THE ALPHA-AMYLASE GENE  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN  
; STREET: 1100 NEW YORK AVENUE, N.W.  
; CITY: WASHINGTON, D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/894,212A  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: KOKULIS, PAUL N.  
; REGISTRATION NUMBER: 16773  
; REFERENCE/DOCKET NUMBER: 95469/C-1195  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 861-3000  
; TELEFAX: (202) 822-0944  
; TELEX: 6714627 CUSH  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 647 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-07-894-212A-8

Query Match 7.5%; Score 140.5; DB 1; Length 647;  
Best Local Similarity 22.8%; Pred. No. 9,3e-05;  
Matches 89; Conservative 60; Mismatches 127; Indels 115; Gaps 23;  
Qy 25 EKAVIPVETLKEIP--FGNITGTYKLKP--KDIIDLVKGGAISDLIELIGTSY 78  
Db 27 EKAVIPVETL--EYEVNPKVIAHTSGPLIEMLODNRFPYDILRSVLRKGQVEIVWAGF 84  
Qy 79 THAIPPLPLSRVEAQDREVK-EFLEFVSPKGFWLPFLAYDPTIPAIKLDNGYEL 136  
Db 85 YEVPLASIP--KEDRERQIRLMKEMAKWSIGDARGWILTERWQPLVKTKEGSDYV 141  
Qy 137 FAMEAMFSAHNSAIKPKIOLYPIKIIKAQKRRYIVSILG 187  
Db 142 IVD-----DYMMSAGISKEEYWPYTEDGCBVIAVFPIDEKL--YLIPFPFWVKY 192

RESULT 4  
US-07-894-212A-2  
Sequence 2, Application US/07894212A  
; Patent No. 536683  
GENERAL INFORMATION:  
; APPLICANT: ASADA, KIYOSO  
; APPLICANT: UEMORI, TAKASHI  
; APPLICANT: MURAKI, HIROYUKI  
; APPLICANT: KATO, IKUNOSHIN  
; APPLICANT: LADERMAN, KENNETH  
; APPLICANT: ANFINSSEN, CHRISTIAN  
TITLE OF INVENTION: THE ALPHA-AMYLASE GENE  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN  
; STREET: 1100 NEW YORK AVENUE, N.W.  
; CITY: WASHINGTON, D.C.  
; ZIP: 20005  
COUNTRY: U.S.A.  
COMPUTER READABLE FORM:  
; MEDIUM TYPE: Disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/894,212A  
; FILING DATE:  
; CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
; NAME: KOKULIS, PAUL N.  
; REGISTRATION NUMBER: 16773  
; REFERENCE/DOCKET NUMBER: 95469/C-1195  
TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 861-3000  
; TELEFAX: (202) 822-0944  
; TELEX: 671427 CUSH  
; INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
; LENGTH: 649 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-07-894-212A-2  
Query Match 7.5%; Score 140.5; DB 1; Length 649;  
Best Local Similarity 22.8%; Pred. No. 9.4e-06;  
Matches 89; Conservative 60; Mismatches 127; Indels 115; Gaps 23;  
Qy 25 EKAYIVIETIILKEEIP--FGLNTGTYLKLPL--KDIIDLVKGGIASDLIRIGTSY 78  
Db 29 EKCYWPFLETL--EELYPNMKVIAHTSGPLWLDNRPYEDPLILASLVKGGQVEIWVAGF 86  
Qy 79 THAILPLPLSRVEAQVQDRBEKE--ELFVSPKGFWLPLPFLAYDPLIPATLKDNGYEL 136  
Db 87 YEPVLASIP--KEDRIQIRLMKEWAKSIGFPDARGWMLTERWQPELVKLIKESIDYV 143

RESULT 5  
US-07-893-928A-1  
Sequence 1, Application US/07893928A  
; Patent No. 5578479  
GENERAL INFORMATION:  
; APPLICANT: LADERMAN, KENNETH  
; APPLICANT: ANFINSSEN, CHRISTIAN  
TITLE OF INVENTION: a-AMYLASE FROM HYPERTHERMOPHILIC  
NUMBER OF INVENTION: ARCHAEBACTERIUM  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN  
; STREET: 1100 NEW YORK AVENUE, N.W.  
; CITY: WASHINGTON, D.C.  
; ZIP: 20005  
COUNTRY: U.S.A.  
COMPUTER READABLE FORM:  
; MEDIUM TYPE: Tape  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/893,928A  
; FILING DATE:  
; CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
; NAME: KOKULIS, PAUL N.  
; REGISTRATION NUMBER: 16773  
; REFERENCE/DOCKET NUMBER: 95470/C-1197  
TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 861-3000  
; TELEFAX: (202) 822-0944  
; TELEX: 671427 CUSH  
; INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
; LENGTH: 650 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-07-893-928A-1  
Query Match 7.5%; Score 140.5; DB 1; Length 650;  
Best Local Similarity 22.8%; Pred. No. 9.4e-06;  
Matches 89; Conservative 60; Mismatches 127; Indels 115; Gaps 23;  
Qy 25 EKAYIVIETIILKEEIP--FGLNTGTYLKLPL--KDIIDLVKGGIASDLIRIGTSY 78  
Db 29 EKCYWPFLETL--EELYPNMKVIAHTSGPLWLDNRPYEDPLILASLVKGGQVEIWVAGF 86  
Qy 79 THAILPLPLSRVEAQVQDRBEKE--ELFVSPKGFWLPLPFLAYDPLIPATLKDNGYEL 136  
Db 87 YEPVLASIP--KEDRIQIRLMKEWAKSIGFPDARGWMLTERWQPELVKLIKESIDYV 143

RESULT 6  
US-09-107-532A-5554  
Sequence 5554, Application US/09107532A  
Patent No. 6583275

GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02454

COMPUTER READABLE FORM:

MEDIUM TYPE: CD/ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09107,532A  
FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/085,598  
FILING DATE: 14-May-1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Arinello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012

TELECOMMUNICATION INFORMATION:

TELEPHONE: (781)893-5907  
TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 5554:

SEQUENCER CHARACTERISTICS:

LENGTH: 227 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES

ORIGINAL SOURCE:

ORGANISM: Enterococcus faecium  
FEATURE:

NAME/KEY: misc feature  
LOCATION: (B) LOCATION 1...227  
SEQUENCE DESCRIPTION: SEQ ID NO: 5554:

RESULT 7  
US-09-710-279-2150  
Sequence 2150, Application US/09710279  
Patent No. 6703492

GENERAL INFORMATION:

APPLICANT: KIMMERLY, WILLIAM JOHN  
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
FILE REFERENCE: PUS480S

CURRENT APPLICATION NUMBER: US/09/710,279

PRIOR APPLICATION NUMBER: 60/164,258  
PRIOR FILING DATE: 1995-11-09  
NUMBER OF SEQ ID NOS: 4472  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO: 2150

LENGTH: 360

TYPE: PRT

ORGANISM: Artificial sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: synthetic

US-09-710-279-2150

Query Ma:ch 5.4%; Score 101.5; DB 4; Length 360;  
Best Local Similarity 19.1%; Pred. No. 0 051; Mismatches 117; Indels 73; Gaps 12;  
Matches 57; Conservative

QY 4 LIVFHGNLOYAEPKSEPKVVKAYKIVIETLKEEPFGLNITGYTKFLKDIILVK 63  
Db 61 LIFPKG-----VKKIVDGGYGSIRKQNNI--NLALHTNDVNPVGUNMLA 107  
QY 64 GGTASDLTEITGP--SYTHAILPLPLSRVQAQDREVKELFEVSPKGFWLPELAYD 121  
Db .08 DQGLENTSMINTNSSTYKVQTPKVKIE-----DFKSLNEL----- 147  
QY .22 PIPAILKQNGYELFADGEAMLSAHNSA-KPKIPLVPHIKAQREKRPVYISVGL 181  
Db 148 ---GLAKEGNTYCFEFSEG-----KGOKPKVGDASPYIGLDS-----IEYV--- 187  
QY 182 REIRKAIKV-----EGKVTIKAVKDIE-AVPV-WWAVNTAVMLGIGRPLMNPCKV 232  
Db 188 ---DEIKLFPKMDKNELEITRRAILDHPSTPVFDPKIKMKSEYGLGICQLNQMT 243  
QY 233 ASWIEDKONILYNGTDEBFIGYRDIAGYRMSVEGLLEV-----IDEINSELCLPSELK 286  
Db 244 LDIFSEYKQKQNPISPTYGTGDPSPKVKVATIGGGGIGPEVKASQGADVFGDVK 302

RESULT 8  
US-09-710-279-2106 Application US/09710279  
; Sequence 2106, Application US/09710279  
; Patent No. 6703432  
; GENERAL INFORMATION:  
; APPLICANT: KIMMERY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PU3480US  
; CURRENT APPLICATION NUMBER: US/09/710,279  
; CURRENT FILING DATE: 2000-11-09  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIORITY FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 2106  
; LENGTH: 376  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; US-09-710-279-2106  
Query Match 5.4%; Score 101.5; DB 4; Length 376;  
Best Local Similarity 19.1%; Pred. No. 0.054; Mismatches 57; Indels 52; Gaps 17; Matches 57;  
Conservative 52; Mismatches 117; Indels 73; Gaps 12;  
QY 4 LVFHENLQYEVIPKSBIPKVEKAYIPVIETLKEIPFGMLNITGYTLKELPKDIDLVK 63  
Db 47 LIFKG-----VKRIVEDGYSITRKLIQNNI-NILALHNDNPKGVNRLMA 93  
QY 64 GGIASDLEIIGT-SYTHAILPLPLSRAEQVQDREYEVBLFEVSPKGFWLPBELAYD 121  
Db 94 DQIGENISMINTNSYYKQTPKPNYIE-----DFDSLNL-----DFDSLNL----- 133  
QY 122 PIPAIPLKONGEYEVIPADGENMLFSAHLSNAIKPKLYPHILIAKOREKRFRTYSYLLGL 181  
Db 134 ---GLAKEGMYEYCFPESEG-----KGQFKPGDASPYIGLDS-----IEV----- 173  
QY 182 RELKAIKUVE-----EGKVTLKAVKDI-EAVPV-----WVANTIAMLGIGRLPMLNPKV 232  
Db 174 ---DEIKLMEWMKONELBETKRAILDNHPVTFPVDFIKNKESBYGLIITQGQNTWT 229  
QY 233 ASWIEDKDNILYGDIEFIGYRDIAGYRMSVEGLEV-----IDELNSCLCLPSBLKH 286  
Db 230 LDESEBVAKQKQINISVRYQHQDPIKKVALLIGSGIGFYYKASQGADWVFDIKH 288  
RESULT 9  
US-08-210-013B-2  
; Sequence 2, Application US/08270013B  
; GENERAL INFORMATION:  
; APPLICANT: Sogabe et al.  
; TITLE OF INVENTION: PROTEIN HAVING HEAT-RESISTANT MALATE  
; TITLE OF INVENTION: DEHYDROGENASE ACTIVITY  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Leydig, Voit & Mayer, Ltd.  
; STREET: Two Prudential Plaza, Suite 4900  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 61601-6780  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/838,418  
; FILING DATE: 17-MAR-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 164701/1993  
; FILING DATE: 02-JUL-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hoover, Allen E.  
; REGISTRATION NUMBER: 37354  
; FILING DATE: 01-JUL-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 164701/1993  
; FILING DATE: 02-JUL-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Green, Robert F.  
; REGISTRATION NUMBER: 27555  
; REFERENCE/DOCKET NUMBER: 63321  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 616-5600  
; TELEFAX: (251) 3533  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 329 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-270-013B-2  
Query Match 5.4%; Score 101; DB 1; Length 329;  
Best Local Similarity 25.3%; Pred. No. 0.05; Mismatches 50; Indels 40; Gaps 7; Matches 39; Conservative 25; Mismatches 50; Indels 40; Gaps 7;  
QY 10 LOYABEIPKSBIPK-----VIEKAYIPVIETLKEIPFGMLNITGYTLKELPKDIDLVK 61  
Db 130 MIVTVFKESGFPKNRIVGQSCVLDTAR--FRTEVABE-----LNIS-----VKDVTGF 175  
QY 62 VKGGIASDLEIIGT-----LPLPLSRAEQVQDREYEVBLFEVSPKGFWLPBELAYD 117  
Db 176 VLGGHEDDMVPLVRYAGGPLEKIPKORLDAVTRTRKGGEIVNLGLGSAYA 235  
QY 118 LAYDPIIPLKDN-----GDEYL 137  
Db 235 ASLVEVNEHAIKQDQRLIPRATVAYLFGYEGIY 269

REFERENCE DOCKET NUMBER: 78339  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (312) 616-5600  
 TELEFAX: (25) 3533  
 INFORMATION FOR SEQ ID NO: 2:  
 LENGTH: 329 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-838-418-2

Query Match 5.4%; Score 101; DB 1; Length 329;  
 Best Local Similarity 25.3%; Pred. No. 0.05; Matches 39; Conservative 25; Mismatches 50; Indels 40; Gaps 7;

Qy 10 LOYAEIPKSEBIPK-----VIEKAYIPIVETIKEBESIPFGUNITGYTUKFLPKPDIIDL 61  
 Db 130 MYYTIVPKESGFPKRNVRIGSVLDTAR--FRTVABE---INIS-----VRDVIGP 175  
 Qy 62 VKGIGIASDLEIIGTSYTHAI-LLPLPLSRSVREQVQRREVKELFEPSPKG--FWLPE 117  
 Db 176 VLGCHGDDMVPLVRSYAGGIPLEKPLKDKRDLAVERTRKGCCBIVNLIGNSGAYAPA 235  
 Qy 118 LAYDPIIAILKON-----GYEYL 137  
 Db 236 ASLVEVMVEAILKQDRRILPAIAVIEGBYEGIY 269

RESULT 11  
 US-09-540-236-318  
 Sequence 3128, Application US/09540236  
 Patent No. 6,673,910  
 GENERAL INFORMATION:  
 APPLICANT: GATC L. Breton et al.  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR  
 TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
 FILE REFERENCE: 27/09-0001  
 CURRENT APPLICATION NUMBER: US/09/540,236  
 CURRENT FILING DATE: 2000-04-04  
 NUMBER OF SEQ ID NOS: 3840  
 SEQ ID NO 3128  
 LENGTH: 653  
 TYPE: PRT  
 ORGANISM: *M. catarrhalis*  
 US-09-540-236-318

Query Match 5.0%; Score 94.5; DB 3; Length 367;  
 Best Local Similarity 18.7%; Pred. No. 0.29; Matches 56; Conservative 52; Mismatches 118; Indels 73; Gaps 12;

Qy 4 LVFHGHLQYAIKPEIPKSEIPKVKYIPIVETIKEBPFGLUNITGYTUKFLPKPDIIDL 63  
 Db 68 LIFKGG-----KTRIVEDGYSITRKLUONI--NLTAHTNLDVNPKGVRNMLA 114  
 Qy 64 GGIASDLEIIGT--SYTHAI-LLPLPLSRSVREQVQRREVKELFEPSPKGFWLPELAYD 121  
 Db 115 DOIGLEMISMINTNSSYKQTFIPRNYG-----DFKOSNLNL----- 154  
 Qy 122 PIPIAIKDNQSYEYLRADGEAMLSAHLNSAIPKIPLYPHLTKAOREBKRFRYISYLU 181  
 Db 155 ---GLAKEGEYNECYCFPESG-----KGQFGPVGADSPYICKLDS-----IEYV-- 194  
 Qy 182 RELRKAIKLVF-----EGKVTKAVKODIE-AVPU--WVAVTHTAVMIGIGRLLMPKV 232  
 Db 195 ---DEBKLEMVKIDNQLEITKRAILDNHPWPMVPEFIMKESBYGLIGLICQLOQMT 250  
 Qy 233 ASWIEDKUNILYGYTDIEFIGVRIADIAGYRMVSEGLEY----IDEUNSELCPSELKH 286  
 Db 251 LDFPSEYAKKQNLNIPSPVRYTGHDSPKPKVAVIGGSSIGEFKASQQLGADVPTGDKH 309

Query Match 5.2%; Score 97; DB 4; Length 653;  
 Best Local Similarity 21.9%; Pred. No. 0.4; Matches 72; Conservative 43; Mismatches 128; Indels 86; Gaps 13;

Qy 22 KVIEKAYIPIVETIKEBESIPFGUNITGYTUKFLPK-----KDIIDLVKGGTAS-- 68  
 Db 95 KRIEPAVKQVITD-REGAPLAANAPLIVFDDYAYABYRUDNEIKTKSETAKOKA 152  
 Qy 69 -----DLEIIGTGSYTHAI-LLPLPLSRSVREAOYD-----REVKELEFEPSPKG 113  
 Db 153 LKKKLEMVLRL-----AAAANYPLEKLEAVGIDHTLTSNSQVKEL--PKGA 201  
 Qy 114 WPELPA-YDPIIAILKONGYEVIFADGEAMLSAHLNSAIPKIPLYPHLKAQREKRP 172  
 Db 202 SSRLVLVILNRSRVAKSVTDLGFLAIGREQYFORYYLQA-EPNQOLQYMAQSDDETQ 260  
 Qy 173 RYISYLLGLRELRAKIKUFLVFEKGVLTKAVKDEIAWPVWAVNTAWMLGIGRLPLANKKK 232  
 Db 261 GYIG-----RAGEAKNERLAGDKD-----VOLRGTRQPOQIEQI 300

RESULT 13  
 US-09-710-1840  
 Sequence 1840, Application US/09710279  
 Patent No. 6,703,492  
 GENERAL INFORMATION:  
 APPLICANT: KIMMELLY, WILLIAM JOHN  
 TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
 FILE REFERENCE: PUS480US  
 CURRENT APPLICATION NUMBER: US/09/710,279  
 CURRENT FILING DATE: 2000-11-09  
 PRIOR APPLICATION NUMBER: 60/164,258  
 PRIOR FILING DATE: 1999-11-09  
 NUMBER OF SEQ ID NOS: 4472  
 SOFTWARE: Patentin Ver. 2.1  
 SEQ ID NO 1840  
 LENGTH: 375  
 TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE: Other  
 OTHER INFORMATION: Description of Artificial Sequence: synthetic  
 OTHER INFORMATION: amino acid sequence  
 US-09-710-1840

Qy 233 ASWIEDKUNILYGYTDIEFIGVRIADIAGYRMVSEGLEY----IDEUNSELCPSELKH 292  
 Qy 301 EPLIEGENIRLTDISRLQVLYKE-----LEOVGRQ-----SARSSGMVVD 343  
 Qy 293 LRT-----SSWAPDKSLRIWREDEGNAR 315

Query Match 4.9%; Score 92; DB 4; Length 375;  
 Best Local Similarity 19.3%; Pred. No. 0.55; Mismatches 72; Conservative 55; Mismatches 109; Indels 138; Gaps 16; Matches 72;

Db 330 VHQQLTEEEEDQFNFLWQDIEISDHMLRAPKLPPTNEESNRPBYYLTEREKS 389  
 Qy 234 SWE---DKONTLUGTDIRFIGYDITAGYRMSTEGGLEVIEDEUNSELCPSELKGR 289  
 Db 390 KWLQESPFDRENFL---PKQNSLRQVPGIQDSVERP---ERSLDLYAPRVRH- 440  
 Qy 3 IDIESG---DFIAFGTSSGKTTALMRMNRMSSTEGETITDGKNEKNDNPVLRSSI 58  
 Db 290 ELYRPSWAD---KSLR-----IMREDEGMRNMLSYNNGELAFLAENS--- 335  
 Qy 112 GFWLPLBLAYD-----ITPAIK-----DNGYEV----- 136  
 Db 441 KLNIDPDSLPLSPKDRLPPPIRCSTIVEGHTG--KIRTISIDPG--LWLATGSDG 496  
 Qy 59 GIVVQIGLQMPHMTKVNIVVPUKJWMSQKDEKAKLILRVUDPBEEDRYFSLG 118  
 Db 336 -ARGWFLPERLDAFRAIMDWRGN 361  
 Qy 137 ------FADSEAMLSSAHNSAIKPKP-KPLYPHLIK---AQREKFRYIYLLG 181  
 Db 497 SVRIVELLTGROVYKQQLINKEINNED 523  
 Db 119 GQQQIGVVRALAAEODILMDEPGALDPTRTDQLVQKQOQGKTEIFVH--- 174  
 Qy 182 RELKAIKLU-----VFECKVTKAVD 203  
 Db 175 -DMDEAIKLAQKICIMTNGOVIQYDTPDNTIRSPANDFVRFQGONRLQDPRNTRVKD 233  
 Qy 204 1EA9PVAWNTAVMIGIRPLMNPKVASSWIEDKNDLILYXGTDIEFIGYRDIAGYRMS 263  
 Db 234 AMIKPTVHRSLSLDAVN---IMBKRV-----DTIVGNDHBLQYLDIDINEC 283  
 Qy 264 VEGGLVBDISNSELSC---LPSLKHSGREYLRTSSWAP---DKSL----- 304  
 Db 284 LRRHKBLDTMQRDITYVRVDRSKLQDSVRTILKRNVRNVPVVDSDNKTLLGLVTRANLVD 343  
 Qy 305 ---RWRRE-DEGN 313  
 Db 344 IVYDSIWELESGN 357

RESULT 14

Sequence 15394, Application US/09248796A

Patent No. 6747137

GENERAL INFORMATION:

APPLICANT: Keith Weinstock et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN

FILE REFERENCE: 107196.132

CURRENT APPLICATION NUMBER: US/09/248, 796A

PRIOR APPLICATION NUMBER: US 60/074, 725

PRIOR FILING DATE: 1998-02-13

PRIOR APPLICATION NUMBER: US 60/094, 409

PRIOR FILING DATE: 1998-08-13

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 31448

LENGTH: 741

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-31448

Query Match 4.9%; Score 91.5; DB 4; Length 741;  
 Best Local Similarity 19.5%; Pred. No. 1.9; Mismatches 84; Conservative 63; Mismatches 131; Indels 153; Gaps 22; Matches 84;

Db 34 TLKKEBIPFGLNITGVTKLKPLPKDIDLVKGIGASDLEIITGTSVTHAI-----LPLLP 87  
 Qy 34 TLKKEBIPFGLNITGVTKLKPLPKDIDLVKGIGASDLEIITGTSVTHAI-----LPLLP 87  
 Db 257 TMMKSHPI---VFGHAVSYVYKDFD- KNG---QFELBLGVNPNNGISSVYDVKKSLP 308  
 Qy 88 LSRVEAQVQDREBEVKELEFV-----SPKG---FWLPL-ELAYDIIIPATLKDNG 132  
 Db 309 ASQEE-----EILHDIEVYSHRPEMAMDSVKGITNLHIPSIVDASMAMIRNSG 361  
 Qy 133 YEVLFDG-----EAMFSAHNSAIKPKP-KPLYPHLIKAQREKFRY 174  
 Db 362 -QMWGKDGKOKDTKAUWPESTYARIYQEMINPKTINGAFDPTTMGSVPNVGLMACKAEY 420  
 Qy 175 ISYLGRLBURKAIKVFEKGKVLKAV-----KDIAVPPWAVNTA-----VM 218  
 Db 421 GSH-----DKTPEMTAG---TMRVLADGSVLMQHDTETGDIWACQPKAPRDWVK 471  
 Qy 219 LGIGR-----LPLMN-----KCVASWEDKNDLILYXGTDIEFIGYRDIAGYR 261  
 Db 472 LAVTRARQSDTPATFWLDPERAHDLERLKVKVLYKHD---LTLGLDLSIMGYNE-AIR 526  
 Qy 262 MSVEGVLLEVDELN-----SEL-----CL 280  
 Db 527 VSMERLIRGKDTISVTVNLVDRYDYLTDLFPIMELGTSAAQKLSIVPLMAGGMYETGAGGS 586  
 Qy 281 PSELKHSGREYLRTSSWADPKSCLRIRED---EGNARLAMLSYMRGELAFLAENSDA 336  
 Db 587 PKHVGQOLVNEENLYRNLDSLBFALAVSLSBTGKTNKAKLKGKALDABTGKLDNKS 646  
 Qy 337 RGWEPLPERL 347  
 Db 647 -----PSRKV 651

Query Match 4.9%; Score 92; DB 4; Length 375;  
 Best Local Similarity 20.2%; Pred. No. 2.1; Length 845;  
 Matches 78; Conservative 64; Mismatches 117; Indels 128; Gaps 21;

Db 38 EIP-FGUNITGVTKLKPLPKDIDLVKGIGASDLEIITGTSVTHAI-----GTSYTHAI 82  
 Qy 202 DEMPHIGYDING-----KRMREPAKGSALDQLESIDLEPEGWGTLQDNTGTS----- 249  
 Qy 83 IPLPLPLSRVQAQVQDREVEKELEF-----VSPKGFWLPLBLAYD-----ITPAIKDONGYEVLFADGE 141  
 Db 250 --LKLTDSEBLIRKIQQQENTDENINP-----VFLFLDWTKD----- 286  
 Qy 142 AMLFSAHNSAIKPKP-KPLYPHLIKAQREKFRYTSYLSLGRERLKAJLUVFEGKV----- 196  
 Db 287 -----EEIMP-----VTAPEPKRKFVPSKHEARVMKJVKARREGRIPNK 329  
 Qy 197 -----TLKAVVKDIEAVPVW---VAINTAVM-LGIGRLP-----LMPKKA 233

Tue Mar 8 09:44:21 2005

us-09-619-032a-4.rai

Page 8

Search completed: March 7, 2005, 21:45:20  
Job time : 45 secB

Copyright (C) 1993 - 2005 Compugen Ltd. GenCore version 5.1.5

OM protein - protein search, using sw model  
Run on: March 7, 2005, 21:42:11 ; Search time 132 Seconds  
Title: US-009-619-032A-4  
Perfect score: 1879  
Sequence: IRLALVPHGHLQYKARIPKSEI. .... RLDLDRFRAIYNDRNGENGSP 907.366 Million cell updates/sec

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1391452 seqs, 329044822 residues

Total number of hits satisfying chosen parameters: 1391452

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA: \*  
1: /cgn2\_6/ptodata/1/pupbaa/US07\_PUBCOMB.pep: \*  
2: /cgn2\_6/ptodata/1/pupbaa/PCT\_NEW\_PUB.pep: \*  
3: /cgn2\_6/ptodata/1/pupbaa/US06\_PUBCOMB.pep: \*  
4: /cgn2\_6/ptodata/1/pupbaa/US05\_PUBCOMB.pep: \*  
5: /cgn2\_6/ptodata/1/pupbaa/US07\_NEW\_PUB.pep: \*  
6: /cgn2\_6/ptodata/1/pupbaa/PCTUS\_PUBCOMB.pep: \*  
7: /cgn2\_6/ptodata/1/pupbaa/US08\_NEW\_PUB.pep: \*  
8: /cgn2\_6/ptodata/1/pupbaa/US08\_PUBCOMB.pep: \*  
9: /cgn2\_6/ptodata/1/pupbaa/US09\_PUBCOMB.pep: \*  
10: /cgn2\_6/ptodata/1/pupbaa/US095\_PUBCOMB.pep: \*  
11: /cgn2\_6/ptodata/1/pupbaa/US09C\_PUBCOMB.pep: \*  
12: /cgn2\_6/ptodata/1/pupbaa/US09\_NEW\_PUB.pep: \*  
13: /cgn2\_6/ptodata/1/pupbaa/US10\_PUBCOMB.pep: \*  
14: /cgn2\_6/ptodata/1/pupbaa/US10C\_PUBCOMB.pep: \*  
15: /cgn2\_6/ptodata/1/pupbaa/US10\_PUBCOMB.pep: \*  
16: /cgn2\_6/ptodata/1/pupbaa/US10\_NEW\_PUB.pep: \*  
17: /cgn2\_6/ptodata/1/pupbaa/US11\_PUB.pep: \*  
18: /cgn2\_6/ptodata/1/pupbaa/US11\_NEW\_PUB.pep: \*  
19: /cgn2\_6/ptodata/1/pupbaa/US60\_PUBCOMB.pep: \*  
20: /cgn2\_6/ptodata/1/pupbaa/US60\_PUBCOMB.pep: \*

Pred. No. 18 is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	1870	99.5	364	9 US-09-886-400-4 Sequence 4, Appli
2	1870	99.5	364	13 US-10-112-357-4 Sequence 4, Appli
3	1870	99.5	364	13 US-10-111-403-4 Sequence 4, Appli
4	1870	99.5	364	13 US-10-116-606-4 Sequence 4, Appli
5	1870	99.5	364	13 US-10-112-331-4 Sequence 4, Appli
6	1870	99.5	364	13 US-10-111-377-4 Sequence 4, Appli
7	1870	99.5	364	13 US-10-116-501-4 Sequence 4, Appli
8	1870	99.5	364	13 US-10-112-442-4 Sequence 4, Appli
9	1870	99.5	364	13 US-10-112-418-4 Sequence 4, Appli
10	1870	99.5	364	13 US-10-114-003-4 Sequence 4, Appli
11	111.5	5.9	890	15 US-10-282-122A-53281 Sequence 53281, A
12	109.5	5.9	312	15 US-10-369-2327 Sequence 23237, A
13	105.5	5.6	573	15 US-10-424-599-175517 Sequence 175517, A

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

No.	Score	Match	Length	DB	ID	Description
1	1870	99.5	364	9	US-09-886-400-4	Sequence 4, Application
2	1870	99.5	364	13	US-10-112-357-4	Sequence 4, Application
3	1870	99.5	364	13	US-10-114-603-4	Sequence 4, Application
4	1870	99.5	364	13	US-11-116-606-4	Sequence 4, Application
5	1870	99.5	364	13	US-10-112-331-4	Sequence 4, Application
6	1870	99.5	364	13	US-10-112-377-4	Sequence 4, Application
7	1870	99.5	364	13	US-10-116-581-4	Sequence 4, Application
8	1870	99.5	364	13	US-10-112-442-4	Sequence 4, Application
9	1870	99.5	364	13	US-10-112-418-4	Sequence 4, Application
10	1870	99.5	364	13	US-10-114-083-4	Sequence 4, Application
11	111.5	5.9	890	15	US-10-282-112A-53281	Sequence 53281, Application
12	105.9	5.8	312	15	US-10-369-493-232517	Sequence 232517, Application
13	105.5	5.6	573	15	US-10-424-599-175517	Sequence 175517, Application

## ALIGNMENTS

Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 7, 2005, 21:42:11 ; Search time 132 Seconds (without alignment@)  
 907.366 Million cell updates/sec

Title: US-09-619-032A-4

Perfect score: 1879

Sequence: 1 LRLALVPHGNLQYAEIPKSEI . . . . . RRLDAFRAIYNDWRGENGEP 364

Scoring table: BL0SUM62

Gapop 10.0 , Gapext 0.5

Searched: 1391452 seqs, 329044822 residues

Total number of hits satisfying chosen parameters: 1391452

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6\_ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*

2: /cgn2\_6\_ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*

3: /cgn2\_6\_ptodata/1/pubpaa/1/ncs2\_new.pdb:\*

4.9	93.9	5.0	2498	16	US-10-495-918-152
44	93.5	5.0	211	16	US-10-437-963-18607
43	93.5	5.0	2049	16	US-10-437-963-19724
42	93.5	5.0	1320	14	US-10-61-051-167
41	93.5	5.0	1320	14	US-10-61-051-167
40	93.5	5.0	1165	15	US-10-282-122A-71765
39	93.5	5.0	993	15	US-10-369-493-20014
38	94.0	5.0	664	14	US-10-032-585-7195
37	94.0	5.0	274	15	US-10-082-122A-5458
36	94.5	5.0	1822	15	US-10-332-443-39
35	94.5	5.0	749	15	US-10-369-493-9015
34	95.0	5.1	2094	16	US-10-437-963-16170
33	95.5	5.1	535	15	US-10-282-122A-5546
32	96.0	5.1	875	15	US-10-282-122A-5580
31	96.0	5.1	530	15	US-10-369-493-10268
30	96.0	5.1	530	15	US-10-202-122A-77126
29	96.0	5.1	798	15	US-10-425-114-37009
28	96.5	5.1	1085	15	US-10-202-122A-5883
27	96.5	5.1	3432	15	US-10-202-122A-6884
26	96.5	5.1	1878	15	US-10-607-631-20
25	97.0	5.2	652	15	US-10-202-122A-63095
24	97.0	5.2	97	5.2	US-10-369-493-55
23	98.0	5.2	510	15	US-10-369-493-55
22	98.5	5.2	813	15	US-10-202-122A-4720
21	100.5	5.3	314	15	US-10-369-493-17395
20	101.0	5.4	441	16	US-10-337-963-12776
19	102.5	5.5	478	15	US-10-696-493-7264
18	102.5	5.5	1150	15	US-10-202-122A-44339
17	103.0	5.5	1147	9	US-09-815-242-12361
16	103.0	5.5	1073	15	US-09-815-242-115544
15	105.5	5.6	1073	15	US-09-815-242-115544

Sequence 55544, A  
Sequence 12361, A  
Sequence 5468, AP  
Sequence 44391, A  
Sequence 4506, AP  
Sequence 7264, AP  
Sequence 12221, A  
Sequence 17395, A  
Sequence 47208, A  
Sequence 56, APP  
Sequence 63055, A  
Sequence 20, APP  
Sequence 37809, A  
Sequence 58833, A  
Sequence 69849, A  
Sequence 10268, A  
Sequence 77165, A  
Sequence 122301, A  
Sequence 52807, A  
Sequence 55460, A  
Sequence 161607, A  
Sequence 9015, AP  
Sequence 39, APP  
Sequence 54582, A  
Sequence 7195, AP  
Sequence 20014, A  
Sequence 71768, A  
Sequence 167, APP  
Sequence 197248, A  
Sequence 16673, A  
Sequence 16071, A  
Sequence 152, APP

Db 61 LVKGGIASDLIEITIGTSYTHAILPLPLPSRVAQVORDEREVHPLSPKGFWLPLAY 120  
 Qy 121 DPTIPAIKLDNGYELFADGEAMLSAHLNSAIKPKIPLYPLIKAREKRYISYLG 180  
 Db 121 DPTIPAIKLDNGYELFADGEAMLSAHLNSAIKPKIPLYPLIKAREKRYISYLG 180  
 Qy 181 LRLRKATKLVFECKVTKAVKDEAVPVWVAINTAVMLGICRPLJMPKKVASHWEDKD 240  
 Db 181 LRLRKATKLVFECKVTKAVKDEAVPVWVAINTAVMLGICRPLJMPKKVASHWEDKD 240  
 Qy 241 NILLYGTDIEFGYRDAGYRMSVEGLIEVIDELNSELCLPSBLKHSGRELYRTSSWAP 300  
 Db 241 NILLYGTDIEFGYRDAGYRMSVEGLIEVIDELNSELCLPSBLKHSGRELYRTSSWAP 300  
 Qy 301 DKSRLIWRDEGENARLNMLSYNMRGELAFLAENS DARGWEPPLPERLDAFRAYNDWRGE 360  
 Db 301 DKSRLIWRDEGENARLNMLSYNMRGELAFLAENS DARGWEPPLPERLDAFRAYNDWRGE 360  
 Qy 361 NGBP 364  
 Db 361 NGBP 364

RESULT 2

US-10-112-357-4  
 ; Sequence 4, Application US/10112357  
 ; Publication No. US20020115099A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: DIVERSA CORPORATION  
 ; APPLICANT: Murphy, Dennis  
 ; APPLICANT: Ried, John  
 ; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND METHODS OF USE THEREOF  
 ; CURRENT FILING DATE: 2002-03-29  
 ; PRIORITY APPLICATION NUMBER: 09/886,400  
 ; PRIORITY FILING DATE: 2001-06-20  
 ; FILE REFERENCE: DIVER1120-4  
 ; CURRENT APPLICATION NUMBER: US/10/112,357  
 ; CURRENT FILING DATE: 2002-03-29  
 ; PRIORITY APPLICATION NUMBER: 09/886,400  
 ; PRIORITY FILING DATE: 1996-03-08  
 ; NUMBER OF SEQ ID NOS: 4  
 ; SOFTWARE: Fast-SEQ for Windows Version 4.0  
 ; SEQ ID NO 4  
 ; LENGTH: 364  
 ; TYPE: PRT  
 ; ORGANISM: Thermococcus alcaliphilus  
 ; US-10-114-403-4

Query Match 99.5%; Score 1870; DB 13; Length 364;  
 Best Local Similarity 99.5%; Pred. No. 4.9e-165; Mismatches 1; Indels 0; Gaps 0;  
 Matches 362; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LRALVFGNQYAEIPKSEIPKVIKEAYIPVIETLKEIPFGLNITGYTUKFLPKOID 60  
 Db 1 LRALVFGNQYAEIPKSEIPKVIKEAYIPVIETLKEIPFGLNITGYTUKFLPKOID 60  
 Qy 61 LVKGGIASDLIEITIGTSYTHAILPLPLPSRVAQVORDEREVHPLSPKGFWLPLAY 120  
 Db 61 LVKGGIASDLIEITIGTSYTHAILPLPLPSRVAQVORDEREVHPLSPKGFWLPLAY 120  
 Qy 121 DPTIPAIKLDNGYELFADGEAMLSAHLNSAIKPKIPLYPLIKAREKRYISYLG 180  
 Db 121 DPTIPAIKLDNGYELFADGEAMLSAHLNSAIKPKIPLYPLIKAREKRYISYLG 180  
 Qy 181 LRLRKATKLVFECKVTKAVKDEAVPVWVAINTAVMLGICRPLJMPKKVASHWEDKD 240  
 Db 181 LRLRKATKLVFECKVTKAVKDEAVPVWVAINTAVMLGICRPLJMPKKVASHWEDKD 240  
 Qy 241 NILLYGTDIEFGYRDAGYRMSVEGLIEVIDELNSELCLPSBLKHSGRELYRTSSWAP 300  
 Db 241 NILLYGTDIEFGYRDAGYRMSVEGLIEVIDELNSELCLPSBLKHSGRELYRTSSWAP 300  
 Qy 361 NGBP 364  
 Db 361 NGBP 364

RESULT 3

US-10-114-403-4  
 ; Sequence 4, Application US/10114403  
 ; Publication No. US20020115100A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: DIVERSA CORPORATION  
 ; APPLICANT: Murphy, Dennis  
 ; APPLICANT: Ried, John  
 ; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND METHODS OF USE THEREOF  
 ; CURRENT FILING DATE: DIVER1120-4  
 ; PRIORITY APPLICATION NUMBER: 09/886,400  
 ; PRIORITY FILING DATE: 2001-06-20  
 ; PRIORITY APPLICATION NUMBER: 09/619,032  
 ; PRIORITY FILING DATE: 2000-07-19  
 ; PRIORITY APPLICATION NUMBER: 09/407,806  
 ; PRIORITY FILING DATE: 1999-09-20  
 ; PRIORITY APPLICATION NUMBER: 09/886,400  
 ; PRIORITY FILING DATE: 1996-03-08  
 ; NUMBER OF SEQ ID NOS: 4  
 ; SOFTWARE: Fast-SEQ for Windows Version 4.0  
 ; SEQ ID NO 4  
 ; LENGTH: 364  
 ; TYPE: PRT  
 ; ORGANISM: Thermococcus alcaliphilus  
 ; US-10-114-403-4

Query Match 99.5%; Score 1870; DB 13; Length 364;  
 Best Local Similarity 99.5%; Pred. No. 4.9e-165; Mismatches 1; Indels 0; Gaps 0;  
 Matches 362; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LRALVFGNQYAEIPKSEIPKVIKEAYIPVIETLKEIPFGLNITGYTUKFLPKOID 60  
 Db 1 LRALVFGNQYAEIPKSEIPKVIKEAYIPVIETLKEIPFGLNITGYTUKFLPKOID 60  
 Qy 61 LVKGGIASDLIEITIGTSYTHAILPLPLPSRVAQVORDEREVHPLSPKGFWLPLAY 120  
 Db 61 LVKGGIASDLIEITIGTSYTHAILPLPLPSRVAQVORDEREVHPLSPKGFWLPLAY 120  
 Qy 121 DPTIPAIKLDNGYELFADGEAMLSAHLNSAIKPKIPLYPLIKAREKRYISYLG 180  
 Db 121 DPTIPAIKLDNGYELFADGEAMLSAHLNSAIKPKIPLYPLIKAREKRYISYLG 180  
 Qy 181 LRLRKATKLVFECKVTKAVKDEAVPVWVAINTAVMLGICRPLJMPKKVASHWEDKD 240  
 Db 181 LRLRKATKLVFECKVTKAVKDEAVPVWVAINTAVMLGICRPLJMPKKVASHWEDKD 240  
 Qy 241 NILLYGTDIEFGYRDAGYRMSVEGLIEVIDELNSELCLPSBLKHSGRELYRTSSWAP 300  
 Db 241 NILLYGTDIEFGYRDAGYRMSVEGLIEVIDELNSELCLPSBLKHSGRELYRTSSWAP 300  
 Qy 361 NGBP 364  
 Db 361 NGBP 364

Query Match 99.5%; Score 1870; DB 13; Length 364;  
 Best Local Similarity 99.5%; Pred. No. 4.9e-165; Mismatches 1; Indels 0; Gaps 0;  
 Matches 362; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LRALVFGNQYAEIPKSEIPKVIKEAYIPVIETLKEIPFGLNITGYTUKFLPKOID 60  
 Db 1 LRALVFGNQYAEIPKSEIPKVIKEAYIPVIETLKEIPFGLNITGYTUKFLPKOID 60  
 Qy 61 LVKGGIASDLIEITIGTSYTHAILPLPLPSRVAQVORDEREVHPLSPKGFWLPLAY 120  
 Db 61 LVKGGIASDLIEITIGTSYTHAILPLPLPSRVAQVORDEREVHPLSPKGFWLPLAY 120  
 Qy 121 DPTIPAIKLDNGYELFADGEAMLSAHLNSAIKPKIPLYPLIKAREKRYISYLG 180  
 Db 121 DPTIPAIKLDNGYELFADGEAMLSAHLNSAIKPKIPLYPLIKAREKRYISYLG 180  
 Qy 181 LRLRKATKLVFECKVTKAVKDEAVPVWVAINTAVMLGICRPLJMPKKVASHWEDKD 240  
 Db 181 LRLRKATKLVFECKVTKAVKDEAVPVWVAINTAVMLGICRPLJMPKKVASHWEDKD 240  
 Qy 241 NILLYGTDIEFGYRDAGYRMSVEGLIEVIDELNSELCLPSBLKHSGRELYRTSSWAP 300  
 Db 241 NILLYGTDIEFGYRDAGYRMSVEGLIEVIDELNSELCLPSBLKHSGRELYRTSSWAP 300

RESULT 4  
 US-10-116-606-4  
 ; Sequence 4, Application US/10116606  
 ; Publication No. US20020119515AI  
 ; GENERAL INFORMATION:  
 ; APPLICANT: DIVERSA CORPORATION  
 ; APPLICANT: Murphy, Dennis  
 ; APPLICANT: Ried, John  
 ; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND METHODS OF USE TH:  
 ; FILE REFERENCE: DIVER1120-4  
 ; CURRENT APPLICATION NUMBER: US/10/116,606  
 ; CURRENT FILING DATE: 2002-04-03  
 ; PRIOR APPLICATION NUMBER: US/09/886,400  
 ; PRIOR FILING DATE: 2001-06-20  
 ; PRIOR APPLICATION NUMBER: 09/619,032  
 ; PRIOR FILING DATE: 2000-07-19  
 ; PRIOR APPLICATION NUMBER: 09/407,806  
 ; PRIOR FILING DATE: 1999-09-20  
 ; PRIOR APPLICATION NUMBER: 08/613,220  
 ; PRIOR FILING DATE: 1996-03-08  
 ; PRIOR APPLICATION NUMBER: 09/407,806  
 ; PRIOR FILING DATE: 2000-07-19  
 ; NUMBER OF SEQ ID NOS: 4  
 ; SOFTWARE: FastSEQ for Windows Version 4.0  
 ; SEQ ID NO: 4  
 ; LENGTH: 364  
 ; TYPE: PRT  
 ; ORGANISM: Thermococcus alcaliphilus  
 US-10-116-606-4

Query Match 99.5%; Score 1870; DB 13; Length 364;  
 Best local Similarity 99.5%; Pred. No. 4.9e-16;  
 Matches 362; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 LRALYVFGHGNLQYAEIPKSEIPKVKIYVITVIETLIKIEIPFGNLTGTYLKFPLKDDID 60  
 Db 1 LRALVFGHGNLQYAEIPKSEIPKVKIYVITVIETLIKIEIPFGNLTGTYLKFPLKDDID 60  
 Qy 61 LVGGGIASDIEIIGTSYTHAILPLPLSRVEAQVDRVEKEELPEVNSKGFWMPLEAY 120  
 Db 61 LVGGGIASDIEIIGTSYTHAILPLPLSRVEAQVDRVEKEELPEVNSKGFWMPLEAY 120  
 Qy 121 DPITPAILKONGYELFLADGECAMLSAHNSAIKPKIPLKPHLKAQREKFRYISYLG 180  
 Db 121 DPITPAILKONGYELFLADGECAMLSAHNSAIKPKIPLKPHLKAQREKFRYISYLG 180  
 Qy 181 LRELRKAIKLVFEGKVTKAVKDIEAVPVWAVTAVMIGRILPMLNPKRVASWIEDK 240  
 Db 181 LRELRKAIKLVFEGKVTKAVKDIEAVPVWAVTAVMIGRILPMLNPKRVASWIEDK 240  
 Qy 241 NILLYGTDIEFIGYRDIAGYRMSVGLLNLVIDEINSELCLPSEKJHSGRELYLRTSSWAP 300  
 Db 241 NILLYGTDIEFIGYRDIAGYRMSVGLLNLVIDEINSELCLPSEKJHSGRELYLRTSSWAP 300  
 Qy 301 DKSLSRIWREDEGNARLNMLSNNMRRGELAFLAENSADRGWELPPLPERRDAFRAINYDWRGE 360  
 Db 301 DKSLSRIWREDEGNARLNMLSNNMRRGELAFLAENSADRGWELPPLPERRDAFRAINYDWRGE 360  
 Qy 361 NGEP 364  
 Db 361 NGEP 364

RESULT 5  
 US-10-112-331-4  
 ; Sequence 4, Application US/10112331  
 ; Publication No. US20020119550AI  
 ; GENERAL INFORMATION:  
 ; APPLICANT: DIVERSA CORPORATION  
 ; APPLICANT: Murphy, Dennis  
 ; APPLICANT: Ried, John  
 ; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND METHODS OF USE TH:  
 ; FILE REFERENCE: DIVER1120-4  
 ; CURRENT APPLICATION NUMBER: US/10/112,331  
 ; CURRENT FILING DATE: 2002-03-29

PRIOR APPLICATION NUMBER: US/09/886,400  
 PRIOR FILING DATE: 2001-05-20  
 PRIOR APPLICATION NUMBER: 09/619,032  
 PRIOR FILING DATE: 2000-07-19  
 PRIOR APPLICATION NUMBER: 09/407,806  
 PRIOR FILING DATE: 1999-09-20  
 PRIOR APPLICATION NUMBER: 08/613,220  
 PRIOR FILING DATE: 1996-03-08  
 NUMBER OF SEQ ID NOS: 4  
 SOFTWARE: FastSEQ for Windows Version 4.0  
 SEQ ID NO: 4  
 LENGTH: 364  
 ; TYPE: PRT  
 ; ORGANISM: Thermococcus alcaliphilus  
 US-10-112-331-4

Query Match 99.5%; Score 1870; DB 13; Length 364;  
 Best local Similarity 99.5%; Pred. No. 4.9e-16;  
 Matches 362; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 LRALYVFGHGNLQYAEIPKSEIPKVKIYVITVIETLIKIEIPFGNLTGTYLKFPLKDDID 60  
 Db 1 LRALYVFGHGNLQYAEIPKSEIPKVKIYVITVIETLIKIEIPFGNLTGTYLKFPLKDDID 60  
 Qy 61 LVGGGIASDIEIIGTSYTHAILPLPLSRVEAQVDRVEKEELPEVNSKGFWMPLEAY 120  
 Db 61 LVGGGIASDIEIIGTSYTHAILPLPLSRVEAQVDRVEKEELPEVNSKGFWMPLEAY 120  
 Qy 121 DPITPAILKONGYELFLADGECAMLSAHNSAIKPKIPLKPHLKAQREKFRYISYLG 180  
 Db 121 DPITPAILKONGYELFLADGECAMLSAHNSAIKPKIPLKPHLKAQREKFRYISYLG 180  
 Qy 181 LRELRKAIKLVFEGKVTKAVKDIEAVPVWAVTAVMIGRILPMLNPKRVASWIEDK 240  
 Db 181 LRELRKAIKLVFEGKVTKAVKDIEAVPVWAVTAVMIGRILPMLNPKRVASWIEDK 240  
 Qy 241 NILLYGTDIEFIGYRDIAGYRMSVGLLNLVIDEINSELCLPSEKJHSGRELYLRTSSWAP 300  
 Db 241 NILLYGTDIEFIGYRDIAGYRMSVGLLNLVIDEINSELCLPSEKJHSGRELYLRTSSWAP 300  
 Qy 301 DKSLSRIWREDEGNARLNMLSNNMRRGELAFLAENSADRGWELPPLPERRDAFRAINYDWRGE 360  
 Db 301 DKSLSRIWREDEGNARLNMLSNNMRRGELAFLAENSADRGWELPPLPERRDAFRAINYDWRGE 360  
 Qy 361 NGEP 364  
 Db 361 NGEP 364

RESULT 6  
 US-10-112-377-4  
 ; Sequence 4, Application US/10112377  
 ; Publication No. US20020120108A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: DIVERSA CORPORATION  
 ; APPLICANT: Murphy, Dennis  
 ; APPLICANT: Ried, John  
 ; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND METHODS OF USE TH:  
 ; FILE REFERENCE: DIVER1120-4  
 ; CURRENT APPLICATION NUMBER: US/10/112,377  
 ; CURRENT FILING DATE: 2002-03-29  
 ; PRIOR APPLICATION NUMBER: 09/886,400  
 ; PRIOR FILING DATE: 2001-06-20  
 ; PRIOR APPLICATION NUMBER: 09/619,032  
 ; PRIOR FILING DATE: 2000-07-19  
 ; PRIOR APPLICATION NUMBER: 09/407,806  
 ; PRIOR FILING DATE: 1999-09-20  
 ; PRIOR APPLICATION NUMBER: 08/613,220  
 ; PRIOR FILING DATE: 1996-03-08  
 ; NUMBER OF SEQ ID NOS: 4  
 ; SOFTWARE: FastSEQ for Windows Version 4.0  
 ; SEQ ID NO: 4  
 ; LENGTH: 364

; TYPE: PRT  
 ; ORGANISM: Thermococcus alcaliphilus  
 US-10-112-377-4  
 Query Match  
 Best Local Similarity 99.5%; Score 1870; DB 13; Length 364;  
 Matches 362; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 LRALVFGNLOQAEIPKEIPKVKIEKAYIPVIETLIKEIPPLGNTGTYLKFLPKOID 60  
 Db 1 LRALVFGNLOQAEIPKEIPKVKIEKAYIPVIETLIKEIPPLGNTGTYLKFLPKOID 60  
 Qy 61 LVKGGLASDLIELIGTSYTHAILPLPLSRVAQVORDREVKEELFESPKGFWLPBLAY 120  
 Db 61 LVKGGLASDLIELIGTSYTHAILPLPLSRVAQVORDREVKEELFESPKGFWLPBLAY 120  
 Qy 121 DPTIPATIKDNGYELYIPADGEAMLFSAHNSAIKPKPLYPHLIKOQREKFRYISYLG 180  
 Db 121 DPTIPATIKDNGYELYIPADGEAMLFSAHNSAIKPKPLYPHLIKOQREKFRYISYLG 180  
 Qy 181 LRLRKAKLKVPEGKVTLKAVDIEAVPVWVAVNTAVMLGGRPLMNPKVVASWEDKD 240  
 Db 181 LRLRKAKLKVPEGKVTLKAVDIEAVPVWVAVNTAVMLGGRPLMNPKVVASWEDKD 240  
 Qy 61 LVKGGLASDLIELIGTSYTHAILPLPLSRVAQVORDREVKEELFESPKGFWLPBLAY 120  
 Db 61 LVKGGLASDLIELIGTSYTHAILPLPLSRVAQVORDREVKEELFESPKGFWLPBLAY 120  
 Qy 121 DPTIPATIKDNGYELYIPADGEAMLFSAHNSAIKPKPLYPHLIKOQREKFRYISYLG 180  
 Db 121 DPTIPATIKDNGYELYIPADGEAMLFSAHNSAIKPKPLYPHLIKOQREKFRYISYLG 180  
 Qy 181 LRLRKAKLKVPEGKVTLKAVDIEAVPVWVAVNTAVMLGGRPLMNPKVVASWEDKD 240  
 Db 181 LRLRKAKLKVPEGKVTLKAVDIEAVPVWVAVNTAVMLGGRPLMNPKVVASWEDKD 240  
 Qy 241 NILLYGTDIEFIGYRDTAGYRMSVEGLIEVIDEINSELCLPSELKHSGRELYRTSSWAP 300  
 Db 241 NILLYGTDIEFIGYRDTAGYRMSVEGLIEVIDEINSELCLPSELKHSGRELYRTSSWAP 300  
 Qy 301 DKSLSRIVREDEBNARLNLSYNNRGEELAFLAENSDARGWEPLPERLDAFRAIYNDWRGE 360  
 Db 301 DKSLSRIVREDEBNARLNLSYNNRGEELAFLAENSDARGWEPLPERLDAFRAIYNDWRGE 360  
 Qy 361 NCERP 364  
 Db 361 NCERP 364  
 Qy 361 NCERP 364  
 Db 361 NCERP 364  
 RESULT 7  
 US-10-116-581-4  
 ; Sequence 4, Application US/10116581  
 ; Publication No. US20020137116A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: DIVERSA CORPORATION  
 ; APPLICANT: Murphy, Dennis  
 ; APPLICANT: Ried, John  
 ; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND  
 ; TITLE OF INVENTION: METHODS OF USE THEREOF  
 ; FILE REFERENCE: DIVER1120-4  
 ; CURRENT APPLICATION NUMBER: US/10/116,581  
 ; CURRENT FILING DATE: 2002-03-29  
 ; PRIORITY APPLICATION NUMBER: 09/886,400  
 ; PRIORITY FILING DATE: 2001-06-20  
 ; PRIORITY APPLICATION NUMBER: 09/619,032  
 ; PRIORITY FILING DATE: 2000-07-19  
 ; PRIORITY APPLICATION NUMBER: 09/407,806  
 ; PRIORITY FILING DATE: 1999-09-20  
 ; PRIORITY APPLICATION NUMBER: 08/613,220  
 ; PRIORITY FILING DATE: 1996-03-08  
 ; NUMBER OF SEQ ID NOS: 4  
 ; SOFTWARE: FastSEQ for Windows Version 4.0  
 ; SEQ ID NO 4  
 ; LENGTH: 364  
 ; TYPE: PRT  
 ; ORGANISM: Thermococcus alcaliphilus  
 US-10-112-442-4  
 Query Match  
 Best Local Similarity 99.5%; Score 1870; DB 13; Length 364;  
 Matches 352; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 LRALVFGNLOQAEIPKEIPKVKIEKAYIPVIETLIKEIPPLGNTGTYLKFLPKOID 60  
 Db 1 LRALVFGNLOQAEIPKEIPKVKIEKAYIPVIETLIKEIPPLGNTGTYLKFLPKOID 60  
 Qy 131 DPTIPATIKDNGYELYIPADGEAMLFSAHNSAIKPKPLYPHLIKOQREKFRYISYLG 180  
 Db 131 DPTIPATIKDNGYELYIPADGEAMLFSAHNSAIKPKPLYPHLIKOQREKFRYISYLG 180  
 Qy 181 LRLRKAKLKVPEGKVTLKAVDIEAVPVWVAVNTAVMLGGRPLMNPKVVASWEDKD 240  
 Db 181 LRLRKAKLKVPEGKVTLKAVDIEAVPVWVAVNTAVMLGGRPLMNPKVVASWEDKD 240  
 Qy 1 LRALVFGNLOQAEIPKEIPKVKIEKAYIPVIETLIKEIPPLGNTGTYLKFLPKOID 60  
 Db 1 LRALVFGNLOQAEIPKEIPKVKIEKAYIPVIETLIKEIPPLGNTGTYLKFLPKOID 60

QY 241 NLLYGTDIRGYDIAGRMSVGGLEVIDESELNSELCPSELKHSGREYLRTSSWAP 300  
 Db 241 NLLYGTDIRGYDIAGRMSVGGLEVIDESELNSELCPSELKHSGREYLRTSSWAP 300  
 QY 301 DKSRLRWEDEBGNARLNLMSNMRCGELAFLAENSARGWPLPERRDAFRAINYDWRGE 360  
 Db 301 DKSRLRWEDEBGNARLNLMSNMRCGELAFLAENSARGWPLPERRDAFRAINYDWRGE 360  
 QY 361 NGEP 364  
 Db 361 NGEP 364

RESULT 9  
 US-10-112-418-4  
 ; Sequence 4, Application US/10112418  
 ; Publication No. US20020155486A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: DIVERSA CORPORATION  
 ; APPLICANT: Murphy, Dennis  
 ; APPLICANT: Ried, John  
 ; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND  
 ; TITLE OF INVENTION: METHODS OF USE THEREOF  
 ; FILE REFERENCE: DIVER1120-4  
 ; CURRENT APPLICATION NUMBER: US/10/112,418  
 ; CURRENT FILING DATE: 2002-04-01  
 ; PRIOR APPLICATION NUMBER: 09/886,400  
 ; PRIOR APPLICATION NUMBER: 09/407,806  
 ; PRIOR FILING DATE: 1996-03-29  
 ; PRIOR FILING DATE: 2000-07-19  
 ; PRIOR FILING DATE: 1999-09-20  
 ; PRIOR FILING DATE: 2001-06-20  
 ; PRIOR APPLICATION NUMBER: 09/619,032  
 ; PRIOR FILING DATE: 2000-07-19  
 ; PRIOR APPLICATION NUMBER: 09/407,806  
 ; PRIOR FILING DATE: 1999-09-20  
 ; PRIOR FILING DATE: 1996-03-08  
 ; NUMBER OF SEQ ID NOS: 4  
 ; SOFTWARE: FastSEQ for Windows Version 4.0  
 ; SBQ ID NO 4  
 ; LENGTH: 364  
 ; TYPE: PRT  
 ; ORGANISM: Thermococcus alcaliphilus  
 US-10-112-418-4

Query Match 99 %; Score 1870; DB 13; Length 364;  
 Best Local Similarity 99 %; Pred. No. 4. 9e-165;  
 Matches 362; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LRALVPHGNIQYAEIPKSEIPKVKVAKYAPIVETIKEEIPFGNLTGTYKFLPKDID 60  
 Db 1 LRALVPHGNIQYAEIPKSEIPKVKVAKYAPIVETIKEEIPFGNLTGTYKFLPKDID 60  
 QY 61 LVGGGASDLEIIGSYTHAILPLPLSVEAQVQDRDVEKEELFESVSPKGFWPLPELAY 120  
 Db 61 LVGGGASDLEIIGSYTHAILPLPLSVEAQVQDRDVEKEELFESVSPKGFWPLPELAY 120  
 QY 121 DPLIPAIKONGYELFADGEAMLSAHIHNSAIKPKIPLYPLIKAQREKRFYISYLG 180  
 Db 121 DPLIPAIKONGYELFADGEAMLSAHIHNSAIKPKIPLYPLIKAQREKRFYISYLG 180  
 QY 61 LVGGGASDLEIIGSYTHAILPLPLSVEAQVQDRDVEKEELFESVSPKGFWPLPELAY 120  
 Db 61 LVGGGASDLEIIGSYTHAILPLPLSVEAQVQDRDVEKEELFESVSPKGFWPLPELAY 120  
 QY 181 LRELRAIKAKUVFEGKVTLKAVKDIAVPPVAVNTAVMIGRLPLNPCKVASHIEDKO 240  
 Db 181 LRELRAIKAKUVFEGKVTLKAVKDIAVPPVAVNTAVMIGRLPLNPCKVASHIEDKO 240  
 QY 241 NLLYGTDIRGYDIAGRMSVGGLEVIDESELNSELCPSELKHSGREYLRTSSWAP 300  
 Db 241 NLLYGTDIRGYDIAGRMSVGGLEVIDESELNSELCPSELKHSGREYLRTSSWAP 300  
 QY 301 DKSRLRWEDEBGNARLNLMSNMRCGELAFLAENSARGWPLPERRDAFRAINYDWRGE 360  
 Db 301 DKSRLRWEDEBGNARLNLMSNMRCGELAFLAENSARGWPLPERRDAFRAINYDWRGE 360  
 QY 361 NGEP 364  
 Db 361 NGEP 364

RESULT 11  
 US-10-282-122A-532A  
 ; Sequence 532A, Application US/10282122A  
 ; Publication No. US20040029129A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Liangsu  
 ; APPLICANT: Zamudio, Carlos  
 ; APPLICANT: Malone, Cheryl

361 NGEP 364

APPLICANT: Haselbeck, Robert  
 APPLICANT: Ohlsen, Kari  
 APPLICANT: Zyskind, Judith  
 APPLICANT: Wall, Daniel  
 APPLICANT: Trawick, John  
 APPLICANT: Carr, Grant  
 APPLICANT: Yamamoto, Robert  
 APPLICANT: Forsyth, R.  
 APPLICANT: Xu, H.  
 TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
 FILE REFERENCE: ELTRA 034A  
 CURRENT FILING DATE: 2003-02-20  
 PRIOR APPLICATION NUMBER: US/10/282,122A  
 PRIOR APPLICATION NUMBER: 60/191,078  
 PRIOR APPLICATION NUMBER: 60/206,848  
 PRIOR FILING DATE: 2000-05-23  
 PRIOR APPLICATION NUMBER: 60/207,727  
 PRIOR FILING DATE: 2000-05-26  
 PRIOR APPLICATION NUMBER: 60/230,335  
 PRIOR APPLICATION NUMBER: 60/230,347  
 PRIOR FILING DATE: 2000-09-09  
 PRIOR APPLICATION NUMBER: 60/242,578  
 PRIOR FILING DATE: 2000-10-23  
 PRIOR APPLICATION NUMBER: 60/253,625  
 PRIOR FILING DATE: 2000-11-27  
 PRIOR APPLICATION NUMBER: 60/257,931  
 PRIOR FILING DATE: 2000-12-22  
 PRIOR APPLICATION NUMBER: 60/267,636  
 PRIOR FILING DATE: 2001-02-09  
 PRIOR APPLICATION NUMBER: 60/269,308  
 PRIOR FILING DATE: 2001-02-16  
 Remaining Prior Application data removed - See File Wrapper or PALM.  
 NUMBER OF SEQ ID NOS: 78614  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 53281  
 LENGTH: 890  
 TYPE: PRT  
 ORGANISM: Clostridium difficile  
 US-10-282-122A-53281

Query Match 5.9%; Score 111.5; DB 15; Length 890;  
 Best Local Similarity 23.6%; Pred. No. 0.49; Mismatches 98; Indels 87; Gaps 16;  
 Matches 72; Conservative 48; Mismatches 98; Indels 87; Gaps 16;

Oy 53 FLPKDIDILVKGGIASDLEIGTSYHAILPLPLSRVEAQVQDRKEVKEELPEVSPKG 112  
 Db 562 RQP-DLIAVGGSAMDGKIMWVYRPEVSPKG 612  
 Oy 113 FWPFLPEAYDPTIP-----ATL--KONGYEVLFADGEAMLFSAHNSAIKPKP 158  
 Db 613 -NGEKAYFAAIPTSAGTGTSEVTPFAVITDQSGVKPLADELMPNMAIDADMMEP 670  
 Oy 159 LPFLH-----IAQREKFRFRYSYL-----IGLREIKRKAIIKLUFR-----GKVTL 198  
 Db 671 -PRLTAASGVDAITALTHALEAVYVSMRLTEPADGLA-----QACKLIFEPYLPRAVNGKNDK 724  
 Oy 199 KAVKDIEAVPVWAVNTA-----VMLGGRPLMN-----P 229  
 Db 725 EREKMANAATWAGMSFANAFGICHLAKHUGAFHIVQHGYANALLIJNEVTKFNCAEP 784  
 Oy 230 KKVASWTEDK-DNLLYGTDEFIGYRDIAGY--RMSVEGLIEVIBLNSELCLPSELK 285  
 Db 785 NKGAFSQYRVDIORYA--EFASFAGIKSQTQEKVNLIKAIDELKAKVGLPKT 841  
 SEQ ID NO L75517  
 LENGTH: 573  
 TYPE: PRT  
 ORGANISM: Glycine max  
 FEATURE: OTHER INFORMATION: Clone ID: PAT\_MRT3847\_12950C.1.pep  
 US-10-424-519-175517

Query Match 5.6%; Score 105.5; DB 15; Length 573;  
 Best Local Similarity 21.3%; Pred. No. 0.95; Mismatches 97; Indels 131; Gaps 19;  
 Matches 77; Conservative 57; Mismatches 97; Indels 131; Gaps 19;

Oy 6 FPNGL-QYAEIRKS-----EPKVIKAYIP-----TETL 35  
 Db 290 FPLSLGKGGVPSSLGLSLGEVPPSPFSPYWCBBPGISTCPSTIAVTQSPNSIETL 349

RESULT 12

US-10-369-493-23237

QY 36 IKEIPFGLNITGYTAKFLPKDIDILVGGIASDLEIIGSYTHAILPLPLPLRVEAQV 95  
 QY 350 --PPPSGASLNLANPLS----VNL----DPVQPLGISMDFP--PFLP---DPLV 388  
 Db 96 QDRDVKEELPEVKRGFWLPELAADP--IPAI-LKDNGTEYLADGEMLFSAHNSA 152  
 QY 389 RMLSLPTSQQ1PTPFTP-----LMDQPIVHPVIVDCSSGCGYLVSGAPMSPS---- 435  
 QY 153 IKPKIPLYPHILKAQREKRFRYIISVILGLRLRKAIKLVFRC-----KGARETRLILUSGSSCQNMQRDTIPAI-- 436  
 Db 436 ---IPPLHPLNVLKPLIPESDAWV----KGARETRLILUSGSSCQNMQRDTIPAI-- 484  
 QY 204 IEAVPVWAVNTAVMLGIGRULPMNPKKVASWIEDKDNLILYGTIEPIGYRDI---- 257  
 Db 485 -----LTPD----ENONNLVAGSRLYTGTRDINIAINS 516  
 QY 258 --AGYRMSVSLGLEYDELNSCLCPSELKHSGRELYLRTSSWAPDK-SLRTRWREDEGNAR 315  
 Db 517 IAAAGLVLVSLGVSVDGVSSELCA----ENYGNLEAVKNS--DGGGAFLDDEGGSS 568  
 QY 316 LN 317  
 Db 569 LD 570

RESULT 14  
 US-10-425-114-55544  
 Sequence 55544, Application US/10425114  
 Publication No. US2004003488A1  
 GENERAL INFORMATION:  
 APPLICANT: Liu, Jingdong  
 APPLICANT: Zhou, Yihua  
 APPLICANT: Kovacic, David K.  
 APPLICANT: Scroen, Steven E  
 APPLICANT: Tabeska, Jack E  
 APPLICANT: Cao, Yongwei  
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 FILE REFERENCE: 38-21(53313)B  
 CURRENT APPLICATION NUMBER: US/10/425,114  
 CURRENT FILING DATE: 2003-04-28  
 NUMBER OF SEQ ID NOS: 73128  
 SEQ ID NO 55544  
 LENGTH: 730  
 TYPE: PRT  
 ORGANISM: Glycine max  
 FEATURE:  
 OTHER INFORMATION: Clone ID: UC-GMFLMINSOY109P02\_FLI.pep  
 US-10-425-114-55544

Query Match 5.6%; Score 105.5; DB 15; Length 730;  
 Best Local Similarity 21.3%; Pred. No. 1.3; Mismatches 77; Conservative 57; Indels 131; Gaps 19;

QY 6 FHGNIT-QYAEIPKS----EIPKVEKAYIPV-----IETL 35  
 Db 447 FLGSIGKLGCVPSIIGLSCVPPISPISPYWCPPGISTCPSSIAVQSPNNSIETL 506  
 QY 36 IKEIPFGLNITGYTAKFLPKDIDILVGGIASDLEIIGSYTHAILPLPLRVEAQV 95  
 Db 507 --PPPSGASLNLANPLS----VNL----DPVQPLGISMDFP--PFLP---DPLV 545  
 QY 96 QDRDVKEELPEVKRGFWLPELAADP--IPAI-LKDNGTEYLADGEMLFSAHNSA 152  
 Db 546 RMLSLPTSQQ1PTPFTP-----LMDQPIVHPVIVDCSSGCGYLVSGAPMSPS---- 592  
 QY 153 IKPKIPLYPHILKAQREKRFRYIISVILGLRLRKAIKLVFRC-----KGARETRLILUSGSSCQNMQRDTIPAI-- 203  
 Db 593 ---PPPLHPLNVLKPLIPESDAWV----KGARETRLILUSGSSCQNMQRDTIPAI-- 641  
 QY 204 IEAVPVWAVNTAVMLGIGRULPMNPKKVASWIEDKDNLILYGTIEPIGYRDI---- 257

RESULT 15  
 US-03-815-242-12361  
 Sequence 12361, Application US/09815242  
 Patent No. US20030051659A1  
 GENERAL INFORMATION:  
 APPLICANT: Hasebeck, Robert  
 APPLICANT: Ohlsen, Kari L.  
 APPLICANT: Zyskind, Judith W.  
 APPLICANT: Wall, Daniel  
 APPLICANT: Travick, John D.  
 APPLICANT: Carr, Grant J.  
 APPLICANT: Yamamoto, Robert T.  
 APPLICANT: Xu, H. Howard  
 TITLE OF INVENTION: Identification of Essential Genes in  
 FILE REFERENCE: ELITA, 01A  
 CURRENT APPLICATION NUMBER: US/09/815,242  
 CURRENT FILING DATE: 2001-03-21  
 PRIOR APPLICATION NUMBER: 60/191,078  
 PRIOR FILING DATE: 2000-03-21  
 PRIOR APPLICATION NUMBER: 60/206,848  
 PRIOR FILING DATE: 2000-05-23  
 PRIOR APPLICATION NUMBER: 60/207,727  
 PRIOR FILING DATE: 2000-05-26  
 PRIOR APPLICATION NUMBER: 60/242,578  
 PRIOR FILING DATE: 2000-10-23  
 PRIOR APPLICATION NUMBER: 60/253,625  
 PRIOR FILING DATE: 2000-11-27  
 PRIOR APPLICATION NUMBER: 60/257,931  
 PRIOR FILING DATE: 2000-12-22  
 PRIOR APPLICATION NUMBER: 60/269,308  
 PRIOR FILING DATE: 2001-02-16  
 NUMBER OF SEQ ID NOS: 14110  
 SOFTWARE: FastSEQ for Windows Version 4.0  
 SEQ ID NO 12361  
 LENGTH: 1073  
 TYPE: PRT  
 ORGANISM: Staphylococcus aureus  
 US-09-815-242-12361

Query Match 5.5%; Score 103; DB 9; Length 1073;  
 Best Local Similarity 21.2%; Pred. No. 4; Mismatches 58; Conservative 48; Indels 78; Gaps 15;

QY 107 FVSKGFWLPELAADPPIPALKONGYEVLFADGEMLFSAHNSAIKPKIPLYPHILKA 166  
 Db 309 EISP-----YDSL---LVKUSTHAISSKQABEKMRSREMRIRGVKNTIPFLINV 357  
 QY 167 DREKER-----RIVSYLGLREKAI----KLVFGKVKTIAVKDIE----- 205  
 Db 358 MNKKEPTSGDYTTKPEETPELFDPSLDRGTKTYLIGNVTINGFPVKEKPKDYLEL 417  
 QY 206 -AVPVWAVNTAVMLGIGRULPMNPKKVASWIEDKDNLILYGTIEPIGYRDI---- 262  
 Db 418 ASIPTVSSKXASGSGTQKQDDEGPKGVABWVKQDVTLL--DTTFRDAHQSLATRV 475  
 QY 263 SVEGLEVIDELNSCLCPSELKHSGRELYLRTSSWAPDK-SLRTRWREDEGNARLMLSY 321  
 Db 476 RTKOMINIAS-----KTAQVKQDFSLMW---GGATPFD-VAY 509  
 QY 322 NMRSGLAFLAENSARGWEPPLERRDAFRATYV 355

Tue Mar 8 09:44:21 2005

us-09-619-032a-4.rapb

Db 510 N-----FLKENP----WERLRLR---KAYPN 529

Search completed: March 7, 2005, 21:56:19  
Job time : 134 secs